Theoretical Computer Science 411 (2010) 1809–1817

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Codes and maximal monoids

Sujin Shin *

Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, Daejeon, 305-701, Republic of Korea

ARTICLE INFO

Article history:
Received 24 April 2009
Received in revised form 29 December 2009
Accepted 29 January 2010
Communicated by D. Perrin

Keywords:
Code
Prefix
Cyclic
Maximal monoid
Indecomposable
Synchronizing

ABSTRACT

Some results are given on maximal monoids in the language of the soficsystem generated by a finite code of words.

1. Introduction

Codes of words correspond to the bases of free submonoids of a free monoid. Their properties have been extensively investigated since the 1950s. The theory of codes is closely related to problems in automata theory, formal languages, and combinatorics on words [4]. This paper mainly concerns the class of sofic systems generated by finite codes of words and maximal monoids. For an irreducible sofic shift X, a set W of words is said to be X-complete if the factors of the monoid W^* form the language of X, denoted $\mathcal{B}(X)$. Complete sets are specially important in automata theory. There is always an X-complete set and there is an X-complete prefix code for any irreducible sofic shift X [5]. In particular, if there is a finite X-complete set, then there is an X-complete prefix code [5]. However, these results do not give any information on whether there is a finite X-complete prefix code. The class of finite complete codes is a more tractable one. The problem to produce X-complete sets leads to the problem of constructing X-indecomposable sets. A set W of words is said to be X-indecomposable if it is the minimal generating sets of W^* and W^* is a maximal monoid in $\mathcal{B}(X)$, i.e., it is maximal under inclusion among the monoids contained in $\mathcal{B}(X)$. Every X-indecomposable set is X-complete and behaves like an "alphabet for $\mathcal{B}(X)$". There exist only finitely many X-indecomposable sets [17]. Various results related to complete sets or indecomposable sets can be also found in [1,2,6,11,13,15,18].

Given any monoid N in $\mathcal{B}(X)$, there is always a maximal monoid M in $\mathcal{B}(X)$ with $M \supseteq N$ [17]. In particular, there is a maximal monoid M in $\mathcal{B}(X)$ containing W^*, so there is an X-indecomposable set V with $V^* \supseteq W^*$. If X is a shift of finite type, then V is finite [16]. In general, it is natural to ask how many finite X-indecomposable sets exist or when W is X-indecomposable.

We will give some answers to these questions. First, in Section 3 we show that, if W is a finite X-indecomposable code, then there is an almost everywhere one-to-one correspondence between infinite paths in the underlying graph induced by W and their labels; that is, the degree of the factor map associated with W is one. The converse holds when W is a

* Tel.: +82 42 350 2713.
E-mail address: IamMarkov@kaist.ac.kr.

0304-3975/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2010.01.035

bifix. In Section 4, we consider the case where W is a finite X-complete prefix code and prove that there is a unique X-indecomposable set $\mathcal{R}$ with $\mathcal{R}^* \supseteq W^*$. Moreover, $\mathcal{R}$ is a finite prefix code. If $\mathcal{R} = W$, i.e., W is X-indecomposable, then it must have a synchronizing word. Finally, in Section 5, for a finite cyclic code W which is X-complete, we provide an upper bound of the number of (finite) X-indecomposable sets V with $V^* \supseteq W^*$. We also present an example of a finite X-complete cyclic code for which such an upper bound cannot be one.

2. Preliminaries

Throughout the work let $\mathcal{A}$ denote a (finite) alphabet. Every word is a word over $\mathcal{A}$ and every shift space is a subshift of the full $\mathcal{A}$-shift $\mathcal{A}^{\mathbb{Z}}$, unless stated otherwise. Let W be a (finite or infinite) set of words. Denote by W^* the free monoid generated by W, i.e., the collection of all finite concatenations of words in W, including the empty word denoted ε. Put $W^+ = W^* \setminus \{\varepsilon\}$. For each $n \geq 1$, define

$$W^n = \{w_1 \cdots w_n | w_i \in W \text{ for } i = 1, \ldots, n\}.$$

In particular, $\mathcal{A}^n$ is the set of all words from $\mathcal{A}^*$ with length n. For each $k \in \mathbb{N}$, let $\mathcal{A}^{k+} = \mathcal{A}^k\mathcal{A}^*$; i.e., $\mathcal{A}^{k+} = \{v \in \mathcal{A}^* || v| \geq k\}$.

A word v is a factor of a word $w \in \mathcal{A}^*$ if there exist $p, s \in \mathcal{A}^*$ with $pvs = w$. If $p = \varepsilon$, i.e., $vs = w$, then v is a prefix of w; if $s = \varepsilon$, i.e., $pv = w$, then v is a suffix of w. Denote by $\mathcal{F}(W)$ the set of all factors of words in W. The set of all prefixes of words in W is denoted $\mathcal{P}(W)$. Similarly, $\mathcal{S}(W)$ is the set of all suffixes of words in W.

Definition. A set W of words is said to be a code or uniquely decipherable if, whenever $u_1 \cdots u_k = v_1 \cdots v_l$ with $u_i, v_j \in W$, then $k = l$ and $u_i = v_i$ for $i = 1, \ldots, k$.

That is, W is a code if every word of W^* admits a unique factorization in words of W, or equivalently, W is the minimal generating set of a free submonoid of $\mathcal{A}^*$.

Definition. A set W of words is said to be a prefix if no word in W is a proper prefix of another word in W. It is said to be a suffix if no word in W is a proper suffix of another word in W. It is said to be a bifix if it is a prefix and a suffix.

Every prefix (or suffix) set is a code. The notion of a prefix code is closely related to that of an automaton. The prefix codes are the easiest to construct and most of the interesting problems on codes can be raised for prefix codes. For works involving prefix codes, see [12,14].

Definition. A set W of words is said to be cyclic if, whenever $p, s \in \mathcal{A}^*$ with $ps \in W$ and v, $svp \in W^*$, we have either $p = \varepsilon$ or $s = \varepsilon$.

Every cyclic set is a code. Cyclic codes define a unique factorization of words written on a circle and they appear in many problems of combinatorics on words. They also have some nice synchronization properties [7,8].

We are interested in finite sets of words, particularly those in these three classes, which are important for various reasons. Each finite set of words corresponds to a labeled graph described as follows. Let W be a finite set of words. Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph with a vertex I and a simple loop labeled w for each $w \in W$ which passes through I and does not intersect any other loop except at I. Here G is the underlying graph of $\mathcal{G}$ and $\mathcal{L}$ the labeling. We call $\mathcal{G}$ the graph presentation for W and I the central vertex of $\mathcal{G}$. Define $\mathsf{X}_W = \mathcal{L}_\infty(\mathsf{X}_G)$, where X_G denotes the edge shift and $\mathcal{L}_\infty$ is the sliding block code induced by $\mathcal{L}$. That is, X_W is the subshift of $\mathcal{A}^{\mathbb{Z}}$ which consists of all possible bi-infinite trips on $\mathcal{G}$. We write $\varphi_W = \mathcal{L}_\infty$. Note that W is X_W-complete; i.e., $\mathcal{F}(W^*) = \mathcal{B}(\mathsf{X}_W)$.

If a shift space X can be described by a (finite) set $\mathcal{F}$ of forbidden words all of which have length $(k+1)$ with $k \geq 0$, i.e., X is the set of sequences in $\mathcal{A}^{\mathbb{Z}}$ which do not contain any words in $\mathcal{F}$, then it is called a k-step shift of finite type. A shift of finite type is a k-step shift of finite type for some $k \geq 0$. A sofic shift is a factor of a shift of finite type, or equivalently, a subshift that can be represented by a labeled graph. A sofic shift is irreducible if it can be represented by a labeled graph $\mathcal{G} = (G, \mathcal{L})$ where G is irreducible; i.e., for every pair of vertices I and J there is a path in G starting at I and terminating at J. Thus every finite set of words generates an irreducible sofic shift, or equivalently, it produces an irreducible labeled graph. (For general background on symbolic dynamics, see [3,10].)

A finite set W of words is said to be simple if any concatenation of more than one words from W does not belong to W. Every code is simple. In general, for a finite set W of words, one can always take a simple subset V of W with $V^* = W^*$, by removing all the words in W that are nontrivial concatenations of words from W. We identify the resulting simple set V with W. For an irreducible sofic shift X, a set W of words is X-indecomposable if and only if it is simple and W^* is a maximal monoid in $\mathcal{B}(X)$. Throughout the work let $\mathcal{W}_\mathcal{A}$ denote the collection of all finite (nonempty) simple sets of words over $\mathcal{A}$.

The following definitions and preliminary results are fundamental to this work.

Definition. Let $W \in \mathcal{W}_\mathcal{A}$. Define

$$W^d = \{w \in \mathcal{A}^* | \mathcal{F}((W \cup \{w\})^*) = \mathcal{F}(W^*)\};$$

that is, a word w is in W^d if $\mathsf{X}_{W\cup\{w\}} = \mathsf{X}_W$. Define

$$W^e = \{w \in \mathcal{A}^* | W^* \cdot w \cdot W^* \subseteq \mathcal{F}(W^*)\};$$

that is, a word w is in W^e if $uwv \in \mathcal{F}(W^*)$ for any $u, v \in W^*$.

Note that $W^* \subseteq W^d \subseteq W^e \subseteq \mathcal{F}(W^*)$. Also, $W^*W^dW^* = W^d$ and $W^*W^eW^* = W^e$. The set W^d is the (finite) union of all maximal monoids in $\mathcal{F}(W^*)$ containing W^* [9].

Proposition 2.1 ([9]). *Let $W \in \mathcal{W}_{\mathcal{A}}$. Then the following are equivalent.*

(i) $W^d = W^e$.
(ii) W^d *is a (maximal) monoid in* $\mathcal{F}(W^*)$.
(iii) W^e *is a (maximal) monoid in* $\mathcal{F}(W^*)$.
(iv) *There is a unique maximal monoid in* $\mathcal{F}(W^*)$ *that contains* W^*.

Thus W is X_W-indecomposable if and only if $W^d = W^*$, or equivalently, $W^e = W^*$.
We introduce two monoids in $\mathcal{F}(W^*)$ which can be regarded as basis elements of W^e.

Definition. Let $W \in \mathcal{W}_{\mathcal{A}}$. Define

$$M_p = \{\alpha \in \mathcal{A}^* | \alpha \cdot W^* \subseteq \mathcal{P}(W^*)\};$$

that is, a word α is in M_p if $\alpha w \in \mathcal{P}(W^*)$ for all $w \in W^*$. Define

$$M_s = \{\beta \in \mathcal{A}^* | W^* \cdot \beta \subseteq \mathcal{S}(W^*)\};$$

that is, a word β is in M_s if $w\beta \in \mathcal{S}(W^*)$ for all $w \in W^*$.

Note that M_p and M_s are monoids in $\mathcal{F}(W^*)$ containing W^*. Also, $M_p \subseteq \mathcal{P}(W^*)$, $M_s \subseteq \mathcal{S}(W^*)$, and $M_s \cdot M_p \subseteq W^e$. These monoids play an important role in characterizing W^d, W^e, or X_W-indecomposable sets, particularly in Sections 4 and 5. The next definition concerns the canonical generators of M_p and M_s.

Definition. Let $W \in \mathcal{W}_{\mathcal{A}}$. Define

$$\mathcal{P}_\circ = \{\alpha \in M_p \setminus \{\varepsilon\} | \alpha \notin W^+ \cdot (M_p \setminus \{\varepsilon\})\};$$

that is, a word α is in $\mathcal{P}_\circ$ if $\alpha \in M_p \setminus \{\varepsilon\}$ and there is no $p \in M_p \setminus \{\varepsilon\}$ with $\alpha \in W^+ \cdot p$. Define

$$\mathcal{S}_\circ = \{\beta \in M_s \setminus \{\varepsilon\} | \beta \notin (M_s \setminus \{\varepsilon\}) \cdot W^+\};$$

that is, a word β is in $\mathcal{S}_\circ$ if $\beta \in M_s \setminus \{\varepsilon\}$ and there is no $s \in M_s \setminus \{\varepsilon\}$ with $\beta \in s \cdot W^+$.

Note that $(\mathcal{P}_\circ)^* = M_p$ and $(\mathcal{S}_\circ)^* = M_s$. One can easily obtain the following result from Proposition 2.1.

Lemma 2.2. *Let $W \in \mathcal{W}_{\mathcal{A}}$. Then the following are equivalent.*

(i) $W^d = W^e = M_p$ ($W^d = W^e = M_s$, *respectively*).
(ii) $W^d \subseteq \mathcal{P}(W^*)$ ($W^d \subseteq \mathcal{S}(W^*)$, *respectively*).
(iii) $W^e \subseteq \mathcal{P}(W^*)$ ($W^e \subseteq \mathcal{S}(W^*)$, *respectively*).

3. Maximal monoids and degree of codes

Let $W \in \mathcal{W}_{\mathcal{A}}$ and $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W with $\varphi = \mathcal{L}_\infty$. Then φ determines how and what the system X_W inherits from the edge shift X_G. Let $w \in \mathcal{F}(W^*) \cap \mathcal{A}^k$ and $1 \le i \le k$. We denote by $d_\varphi(w, i)$ the number of edges e in G such that there is a path $\tau = \tau_1 \cdots \tau_k$ in G with $\mathcal{L}(\tau) = w$ and $\tau_i = e$. The number

$$d_\varphi = \min\{d_\varphi(w, i) | w \in \mathcal{F}(W^*), 1 \le i \le |w|\}$$

is called the degree of φ. Let φ be finite-to-one, or equivalently, let W be a code. Then there is $r \ge 1$ such that $|\mathcal{L}^{-1}(v)| \le r$ for all $v \in \mathcal{F}(W^*)$. The number of pre-images under φ can vary through a finite set of positive integers. The degree of φ is exactly the minimum number of pre-images under φ (see [10]). We show that, if W is X_W-indecomposable, then $d_\varphi = 1$ (see Proposition 3.2).

Lemma 3.1. *Let $W \in \mathcal{W}_{\mathcal{A}}$ be a code and $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W with $\varphi = \mathcal{L}_\infty$. Let $w \in W^+$ and $s = |\mathcal{L}^{-1}(w)|$. Then the following are equivalent.*

(i) *Let $r \ge s + 1$. Then $d_\varphi(w^{2r}, |w^r|) = 1$.*
(ii) *There is $n \in \mathbb{N}$ such that $d_\varphi(w^n, i) = 1$ for some i, $1 \le i \le |w^n|$.*
(iii) *There is a unique cycle in G labeled w^n for each $n \in \mathbb{N}$.*
(iv) *There is a unique cycle in G labeled w^n for each $n = 1, \ldots, s$.*

Proof. Let I be the central vertex of $\mathcal{G}$. Let $k = |w|$ and $\eta = \eta_1 \cdots \eta_k$ denote the cycle in G starting at I labeled w. It is clear that (i) implies (ii), and (iii) implies (iv). To show that (ii) implies (iii), let ζ be a cycle that does not start at I and is labeled w^m for some $m \geq 1$. Since W is a code, it follows that $d_\varphi(w^n, i) \geq 2$ for all $n \in \mathbb{N}$ and all $i = 1, \ldots, |w^n|$. Thus (ii) implies (iii). It remains to show that (iv) implies (i).

Let $d_\varphi(w^{2r}, |w^r|) \geq 2$. Then there is a path

$$\bar{\zeta} = \zeta_1^{(1)} \cdots \zeta_k^{(1)} \cdots \zeta_1^{(r)} \cdots \zeta_k^{(r)} \zeta_1^{(r+1)} \cdots \zeta_k^{(r+1)} \cdots \zeta_1^{(2r)} \cdots \zeta_k^{(2r)}$$

labeled w^{2r} such that $\zeta^{(i)} = \zeta_1^{(i)} \cdots \zeta_k^{(i)} \in \mathcal{L}^{-1}(w)$ for all $i = 1, \ldots, 2r$ and $\zeta_k^{(r)} \neq \eta_k$. Also, there exist i_1, i_2, i_3, i_4 with $1 \leq i_1 < i_2 \leq r < i_3 < i_4 \leq 2r$ such that $\zeta^{(i_1)} = \zeta^{(i_2)}$ and $\zeta^{(i_3)} = \zeta^{(i_4)}$. If $\zeta_1^{(i_1)}$ and $\zeta^{(i_3)}$ both start at I, then $\zeta^{(i_4-1)}$ ends at I. Put

$$\theta = \zeta^{(i_1)} \cdots \zeta^{(i_4-1)} = \theta_1 \cdots \theta_{(i_4-i_1)k}.$$

Then θ is a cycle starting at I such that $\mathcal{L}(\theta) = w^{i_4-i_1}$ and $\theta_{(r-i_1+1)k} = \zeta_k^{(r)} \neq \eta_k$. This is impossible, since W is a code. So either $\zeta_1^{(i_1)}$ or $\zeta^{(i_3)}$ does not start at I. Let

$$\rho = \zeta^{(i_1)} \cdots \zeta^{(i_2-1)} \quad \text{and} \quad \bar{\rho} = \zeta^{(i_3)} \cdots \zeta^{(i_4-1)}.$$

Then ρ and $\bar{\rho}$ are cycles labeled by $w^{i_2-i_1}$ and $w^{i_4-i_3}$, respectively. Also, either ρ or $\bar{\rho}$ does not start at I. Thus there is a cycle ζ that does not start at I and is labeled by w^m for some m, $1 \leq m < r$. Thus (iv) implies (i). This completes the proof. □

We introduce the notion of an intrinsically synchronizing word which is widely used throughout the work. Every irreducible sofic shift X has an intrinsically synchronizing word w in $\mathcal{B}(X)$; i.e., for any $u, v \in \mathcal{B}(X)$ with $uw, wv \in \mathcal{B}(X)$, we have $uwv \in \mathcal{B}(X)$ (see [10]). In particular, if $W \in \mathcal{W}_\mathcal{A}$, then there is an intrinsically synchronizing word for X_W. Denote by $\mathcal{I}_s$ the set of intrinsically synchronizing words for X_W. Any extension of a word in $\mathcal{I}_s$ is in $\mathcal{I}_s$. If X_W is a k-step shift of finite type for some $k \geq 1$, then every word in $\mathcal{F}(W^*) \cap \mathcal{A}^{k+}$ is an intrinsically synchronizing word for X_W.

Proposition 3.2. *Let $W \in \mathcal{W}_\mathcal{A}$ be a code and X_W-indecomposable. Put $\varphi = \varphi_W$. Then there is $r \in \mathbb{N}$ such that $d_\varphi(w^{2r}, |w^r|) = 1$ for all $w \in \mathcal{I}_s \cap W^+$. In particular, the degree of φ is one.*

Proof. Let $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W with the central vertex I. Choose $r \in \mathbb{N}$ so that $|\mathcal{L}^{-1}(v)| < r$ for all $v \in \mathcal{F}(W^*)$. Let $w \in \mathcal{I}_s \cap W^+$ and $k = |w|$. Let $\eta = \eta_1 \cdots \eta_k$ denote the cycle starting at I labeled w. Suppose that $d_\varphi(w^{2r}, |w^r|) \geq 2$. By Lemma 3.1, there is a cycle ζ that does not start at I and is labeled w^n for some $n \geq 1$. Let $d = d_\varphi(w, k)$. Then there exist d edges $e_1, \ldots, e_d$ in G such that, whenever $\tau = \tau_1 \cdots \tau_k$ is a path in G with $\mathcal{L}(\tau) = w$, then $\tau_k \in \{e_1, \ldots, e_d\}$. Let $\pi = \pi_1 \cdots \pi_k$ be the prefix subpath of ζ labeled w; that is, there exist a cycle ξ starting at I and two paths α, γ such that $\zeta = \pi\gamma\xi\alpha$ and $\mathcal{L}(\gamma\xi\alpha) = w^{n-1}$. Put $p = \mathcal{L}(\alpha), s = \mathcal{L}(\gamma)$ and $u = \mathcal{L}(\xi)$. We may assume that $\eta_k = e_1$ and $\pi_k = e_2$.

Let $v = w_1 \cdots w_{k-1}$, so $w = vw_k$. Since $pw \in \mathcal{P}(W^*)$ and $w \in \mathcal{I}_s \cap W^*$, we have $pw \in W^e$; hence $pvw_k = pw \in W^*$. Similarly, $vw_ks = ws \in W^*$. For each $j = 1, \ldots, d$, let K_j be the initial vertex of e_j and J_j the terminal vertex of e_j. Then there exist $l, m \in \{1, \ldots, d\}$ such that $J_l = I$ and

(i) there is a path ρ_l from I to K_l labeled pv;
(ii) there is a path κ_m from I to K_m labeled v;
(iii) there is a path ζ_m from J_m to I labeled s.

Put

$$\tau = \kappa_m e_m \zeta_m \xi \rho_l e_l.$$

Then τ is a cycle starting at I and

$$\mathcal{L}(\tau) = vw_k sup vw_k = w^{n+1} = \mathcal{L}(\eta^{n+1}).$$

Since W is a code, we have $l = m = 1$. So $\rho_1 e_1 \zeta_1$ is a cycle starting at I labeled pws. Meanwhile, $\alpha\pi_1 \cdots \pi_k\gamma$ is a cycle starting at I labeled pws and $\pi_k = e_2$. This is impossible, since W is a code. Thus $d_\varphi(w^{2r}, |w^r|) = 1$. Therefore $d_\varphi = 1$. □

Example 3.1. Let $W = \{1, 00, 1000\}$ and $\varphi = \varphi_W$. Then W is a code but not X_W-indecomposable, since $0^3 \in W^e \setminus W^*$. Let $w = 00$. Then $w \in \mathcal{I}_s \cap W$. But $d_\varphi(w^n, i) \geq 2$ for all $n \geq 1$ and all $i = 1, \ldots, |w^n|$.

Example 3.2. Let $W = \{1, 00, 000\}$ and $\varphi = \varphi_W$. Then W is X_W-indecomposable but not a code. Let $w = 00$. Then $w \in \mathcal{I}_s \cap W$. But $d_\varphi(w^n, i) \geq 2$ for all $n \geq 1$ and all $i = 1, \ldots, |w^n|$.

Lemma 3.3. *Let $W \in \mathcal{W}_\mathcal{A}$ and $\varphi = \varphi_W$. Then the following are equivalent.*

(i) *W is cyclic.*
(ii) *There is $r \in \mathbb{N}$ such that $d_\varphi(w^{2r}, |w^r|) = 1$ for all $w \in W^+$.*
(iii) *There is $r \in \mathbb{N}$ such that $d_\varphi(w^{2r}, |w^r|) = 1$ for all $w \in \mathcal{I}_s \cap W^+$. Also, X_W is a shift of finite type.*

Proof. First, by Lemma 3.1, (i) implies (ii) and hence (iii). Next, if X_W is a shift of finite type, then, given $w \in W^+$, we have $w^n \in \mathcal{I}_s \cap W^+$ for all n large enough. So (iii) implies (ii). Thus it suffices to show that (ii) implies (i). To this end, assume that (ii) holds and that W is not cyclic; that is, φ is not one-to-one. Then φ is not one-to-one on the periodic points of X_W (see [10]). Hence there is a periodic point $x \in \mathsf{X}_W$ that has at least 2 pre-images. Choose $w \in W^+$ so that $x = w^\infty$. Let $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W with the central vertex I. Then there exist two distinct cycles η, τ in G labeled w^l for some $l \geq 1$ such that η starts at I and τ starts at J. We may assume that $l = 1$.

Let $\eta = \eta_1 \cdots \eta_k$ and $\tau = \tau_1 \cdots \tau_k$, where $|w| = k \geq 1$. Let

$$\mathcal{J} = \{i | 1 \leq i \leq k \text{ and } \eta_i = \tau_i\}.$$

If $\mathcal{J} = \emptyset$, then $d_\varphi(w^n, i) \geq 2$ for all $n \geq 1$ and all $i = 1, \ldots, |w^n|$, which is a contradiction. Hence $\mathcal{J} \neq \emptyset$. Put

$$\mathcal{J} = \{i_1 < i_2 < \cdots < i_s\},$$

where $s \geq 1$. Then $\eta_{i_1} (= \tau_{i_1})$ starts at I and $\eta_{i_s} (= \tau_{i_s})$ terminates at I. If either $i_1 > 1$ or $i_s < k$, then let

$$\bar{\eta} = \eta_1 \cdots \eta_{i_1 - 1} \eta_{i_s + 1} \cdots \eta_k, \qquad \bar{\tau} = \tau_1 \cdots \tau_{i_1 - 1} \tau_{i_s + 1} \cdots \tau_k,$$

and

$$\bar{w} = w_1 \cdots w_{i_1 - 1} w_{i_s + 1} \cdots w_k.$$

Note that $\bar{w} \in W^+$ and that $\bar{\eta}$ and $\bar{\tau}$ are two cycles labeled $\bar{w}$. Also, $\eta_i \neq \tau_i$ for all $i = 1, \ldots, i_1 - 1$ and for all $i = i_s + 1, \ldots, k$. It follows that $d_\varphi(\bar{w}^n, i) \geq 2$ for all $n \geq 1$ and all $i = 1, \ldots, |\bar{w}^n|$, which is a contradiction. Thus $i_1 = 1$ and $i_s = k$, so $I = J$.

Let

$$t = \max\{i | 1 \leq i \leq k, i \notin \mathcal{J}\}.$$

Such t exists, since $\eta \neq \tau$. Since $t < k$ and $\eta_{t+1} = \tau_{t+1}$, it follows that η_t and τ_t end at I. Define $\tilde{\eta} = \eta_1 \cdots \eta_t$ and $\tilde{\tau} = \tau_1 \cdots \tau_t$. Let $\tilde{w} = w_1 \cdots w_t$. Then $\tilde{\eta}$ and $\tilde{\tau}$ are two cycles labeled $\tilde{w}$ and $\tilde{w} \in W^+$. Also, $\eta_t \neq \tau_t$. So $d_\varphi(\tilde{w}^{2r}, |\tilde{w}^r|) = 2$ for any $r \geq 1$, which is a contradiction. Thus (ii) implies (i). □

Proposition 3.2 and Lemma 3.3 imply the following result, which appears in [9].

Corollary 3.4. *Let $W \in \mathcal{W}_\mathcal{A}$ be a code and X_W-indecomposable. Then it is cyclic if and only if X_W is a shift of finite type.*

There is an example of a cyclic code $W \in \mathcal{W}_\mathcal{A}$ that is not X_W-indecomposable (see Example 5.1). Thus the converse of Proposition 3.2 does not hold. That is, given a code $W \in \mathcal{W}_\mathcal{A}$, the property that there is $r \in \mathbb{N}$ such that $d_\varphi(w^{2r}, |w^r|) = 1$ for all $w \in \mathcal{I}_s \cap W^+$ with $\varphi = \varphi_W$ does not guarantee that W is X_W-indecomposable.

4. Prefix codes and maximal monoids

In this section we investigate X_W-indecomposable sets for a finite prefix code W. Let $W \in \mathcal{W}_\mathcal{A}$ be a prefix code. We show that there is a unique X_W-indecomposable set $\mathcal{R}$ with $\mathcal{R}^* \supseteq W^*$. Also, $\mathcal{R}$ is a finite prefix code (see Proposition 4.2). If $\mathcal{R} = W$, i.e., W is X_W-indecomposable, then it has a synchronizing word (see Proposition 4.6). The converse holds if W is a bifix.

Lemma 4.1. *Let $W \in \mathcal{W}_\mathcal{A}$ be a prefix. Then $W^e = M_p$.*

Proof. Let $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W with the central vertex I. Let $d = d_\varphi$, where $\varphi = \mathcal{L}_\infty$. Then there exist $w = w_1 \cdots w_k \in \mathcal{F}(W^*)$ and i, $1 \leq i \leq k$, such that $d_\varphi(w, i) = d$. That is, there exist d edges $e_1, \ldots, e_d$ in G such that, whenever $\tau = \tau_1 \cdots \tau_k$ is a path in G with $\mathcal{L}(\tau) = w$, then $\tau_i \in \{e_1, \ldots, e_d\}$. We may assume that $w \in W^*$ and that there is a cycle $\eta = \eta_1 \cdots \eta_k$ in G starting at I with $\mathcal{L}(\eta) = w$ and $\eta_i = e_1$. Let J be the vertex at which e_1 terminates.

Let $\alpha \in W^e$. Then $w\alpha \in W^e$ and $d_\varphi(w\alpha, i) = d$. If $\tau = \tau_1 \cdots \tau_k \tilde{\tau}$ is a path in G such that $\mathcal{L}(\tau_j) = w_j$ for $j = 1, \ldots, k$ and $\mathcal{L}(\tilde{\tau}) = \alpha$, then $\tau_i \in \{e_1, \ldots, e_d\}$. Hence there must be a path $\gamma = \gamma_1 \cdots \gamma_k \tilde{\gamma}$ in G with $\gamma_i = e_1$ such that $\mathcal{L}(\gamma_j) = w_j$ for $j = 1, \ldots, k$ and $\mathcal{L}(\tilde{\gamma}) = \alpha$. Note that the path $\gamma_{i+1} \cdots \gamma_k$ starts at J and

$$\mathcal{L}(\gamma_{i+1} \cdots \gamma_k) = w_{i+1} \cdots w_k \in \mathcal{S}(W^*).$$

It follows that the path $\gamma_1 \cdots \gamma_k$ terminates at I, since W is a prefix. So $\tilde{\gamma}$ starts at I. Hence $\alpha \in \mathcal{P}(W^*)$. Thus $W^e \subseteq \mathcal{P}(W^*)$. Therefore $W^e = M_p$ by Lemma 2.2. □

Remark 4.1. *Let $W \in \mathcal{W}_\mathcal{A}$ be a suffix. Then $W^e = M_s$.*

Let $W \in \mathcal{W}_\mathcal{A}$. Define

$$\mathcal{R} = \{\alpha \in \mathcal{P}_\circ | \{\alpha\}^+ \cap W \neq \emptyset\}.$$

That is, $\alpha \in \mathcal{P}_\circ$ is in $\mathcal{R}$ if and only if $\alpha^k \in W$ for some $k \geq 1$. Note that $\mathcal{R}$ is finite and $W \subseteq \mathcal{R} \subseteq \mathcal{P}_\circ$.

Proposition 4.2. *Let $W \in \mathcal{W}_\mathcal{A}$ be a prefix. Then $W^e = \mathcal{R}^*$. Also, $\mathcal{R}$ (after simplification) is a prefix.*

Proof. First, observe that if $\alpha \in M_p$ and $\alpha = up$ for some $u \in W^*$ and $p \in \mathcal{P}(W^*)$, then $p \in M_p$. This is true because W is a prefix. We claim that $\mathcal{P}_\circ = M_p \cap \mathcal{P}(W)$. To see this, let $\alpha \in \mathcal{P}_\circ \setminus \mathcal{P}(W)$. Then there exist $u \in W^+$ and $p \in \mathcal{P}(W)$ such that $\alpha = up$. If $w \in W^*$, then $\alpha w = upw \in \mathcal{P}(W^*)$. It follows that $pw \in \mathcal{P}(W^*)$, since W is a prefix. Hence $p \in M_p$, which is a contradiction. Thus $\mathcal{P}_\circ \subseteq M_p \cap \mathcal{P}(W)$. If $\alpha \in M_p \cap \mathcal{P}(W)$, then $\alpha \in \mathcal{P}_\circ$, since W is a prefix. Therefore $\mathcal{P}_\circ = M_p \cap \mathcal{P}(W)$.

Next, we will show that $\mathcal{P}_\circ = \mathcal{R}^* \cap \mathcal{P}(W)$. To see this, let $\alpha \in \mathcal{P}_\circ$ be a smallest element in $\mathcal{P}_\circ \setminus \mathcal{R}^*$. Since $\alpha^k \in M_p \setminus W$ for all $k \geq 1$, there exist $b, a \in \mathcal{P}(W^*) \setminus \{\varepsilon\}$, $v \in W^*$ and $k \geq 1$ such that $\alpha = ba$ and $\alpha^k b \in W$. Set $u = \alpha^k b$. Since $\alpha^{k+1} = ua$ and $\alpha u = uab$, it follows from the above claim that $a, ab \in M_p$. Choose $v \in W^*$ and $p \in \mathcal{P}_\circ$ so that $a = vp$. Since $\alpha u = uvpb$, we have $pb \in M_p$. Since $|p| < |\alpha|$, by the assumption, $p \in \mathcal{R}^*$; that is, there exist $q_1, \ldots, q_s \in \mathcal{R}$ and $k_1, \ldots, k_s \in \mathbb{N}, s \geq 1$, such that $p = q_1 \cdots q_s$ and $(q_i)^{k_i} \in W$ for each $i = 1, \ldots, s$. Consider

$$(q_1)^{k_1-1} pb = (q_1)^{k_1-1} q_1 q_2 \cdots q_s b = (q_1)^{k_1} q_2 \cdots q_s b \in M_p.$$

Then $q_2 \cdots q_s b \in M_p$. Next, consider

$$(q_2)^{k_2-1} q_2 q_3 \cdots q_s b = (q_2)^{k_2} q_3 \cdots q_s b \in M_p.$$

Then $q_3 \cdots q_s b \in M_p$. Continuing this process, we obtain $q_s b \in M_p$. Finally, consider

$$(q_s)^{k_s-1} q_s b = (q_s)^{k_s} b \in M_p$$

to obtain $b \in M_p$. Since $\alpha \in \mathcal{P}_\circ$, it follows that $b \in \mathcal{P}_\circ$. So $b \in \mathcal{R}^*$. Since $W^* \subseteq \mathcal{R}^*$, we have $\alpha = bvp \in \mathcal{R}^*$, which is a contradiction. Thus $\mathcal{P}_\circ \subseteq \mathcal{R}^*$. Since $\mathcal{R}^* \subseteq M_p$, it follows that $\mathcal{P}_\circ = \mathcal{R}^* \cap \mathcal{P}(W)$. Thus $W^e = \mathcal{R}^*$, so W^e is finitely generated.

To show that $\mathcal{R}$ (after simplification) is a prefix, let $\alpha \in \mathcal{R}$ and $\beta \in \mathcal{A}^*$ with $\alpha\beta \in \mathcal{R}$. Then there exist $k, l \geq 1$ such that $\alpha^k, (\alpha\beta)^l \in W$. Since $\alpha^k \beta = \alpha^{k-1}(\alpha\beta) \in M_p$, we have $\beta \in M_p$. So $\beta \in \mathcal{R}^*$. It follows that $\beta = \varepsilon$. Thus $\mathcal{R}$ is a prefix. □

Let $W \in \mathcal{W}_\mathcal{A}$. A word $w \in \mathcal{F}(W^*)$ is primitive if there is no $u \in \mathcal{A}^*$ with $u^k = w$ for some $k \geq 2$. We say that W is primitive if every word $w \in W$ is primitive. If W is pure, i.e., for any $v \in \mathcal{A}^* \setminus W^*$, we have $v^k \notin W^*$ for all $k \in \mathbb{N}$, then it is primitive.

Corollary 4.3. *Let $W \in \mathcal{W}_\mathcal{A}$ be a prefix or a suffix. If W is primitive, then it is X_W-indecomposable.*

The following is immediate from Corollary 3.4.

Corollary 4.4. *Let $W \in \mathcal{W}_\mathcal{A}$ be a prefix (or suffix) and X_W be a shift of finite type. Then the following are equivalent.*

(i) *W is cyclic.*
(ii) *W is pure.*
(iii) *W is primitive.*
(iv) *W is X_W-indecomposable.*

As mentioned before, there is an example of a cyclic code $W \in \mathcal{W}_\mathcal{A}$ that is not X_W-indecomposable (see Example 5.1). Every cyclic code is primitive. Thus if W is merely a code (even if W is a cyclic code), then the conclusion of Corollary 4.3 need not be true. Consequently, when W is a code and it generates a shift of finite type, the conclusion of Corollary 4.4 need not hold true.

Let $W \in \mathcal{W}_\mathcal{A}$ and $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W. A synchronizing word for W is a word $w \in \mathcal{F}(W^*)$ such that all paths in G labeled w end up at the same vertex. If w is a synchronizing word for W, then so is vw for any $v \in \mathcal{F}(W^*)$ with $vw \in \mathcal{F}(W^*)$. A synchronizing word is a classical and elementary notion in automata theory. A word $w \in \mathcal{F}(W^*)$ is called an r-synchronizing word for W if all paths in G presenting w start at the same vertex.

Proposition 4.5. *Let $W \in \mathcal{W}_\mathcal{A}$ be a prefix. If there is an r-synchronizing word for W, then W is X_W-indecomposable.*

Proof. Let $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W with the central vertex I. If there is an r-synchronizing word for W, then there is $u \in \mathcal{S}(W^*)$ such that all paths in G presenting u start at the same vertex, say J. Take the shortest path γ from I to J and let $p = \mathcal{L}(\gamma)$. Replacing w with pu, if necessary, we may assume that $w \in W^*$. Since W is a prefix, there is a unique path τ labeled w and τ is a cycle starting at I. If $\alpha \in W^e$, then $\alpha w \in \mathcal{P}(W^*)$, since W is a prefix. It follows that $\alpha \in W^*$. Thus W is X_W-indecomposable. □

Proposition 4.6. *Let $W \in \mathcal{W}_\mathcal{A}$ be a prefix and X_W-indecomposable. Put $\varphi = \varphi_W$.*

(1) *There is $r \in \mathbb{N}$ such that $d_\varphi(w^r, |w^r|) = 1$ for all $w \in \mathcal{U}_s \cap W^+$.*
(2) *Every word in $\mathcal{U}_s \cap \mathcal{S}(W^*)$ is a synchronizing word for W.*

Proof. Let $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W with the central vertex I.

(1) Choose $r \in \mathbb{N}$ so that $|\mathcal{L}^{-1}(v)| < r$ for all $v \in \mathcal{F}(W^*)$. Let $w \in \mathcal{U}_s \cap W^+$ and $k = |w|$. Then $d_\varphi(w^{2r}, |w^r|) = 1$ by Proposition 3.2. Let $\eta = \eta_1 \cdots \eta_k$ denote the cycle starting at I labeled w. If $d_\varphi(w^r, |w^r|) \geq 2$, then there is a path

$$\bar{\zeta} = \zeta_1^{(1)} \cdots \zeta_k^{(1)} \zeta_1^{(2)} \cdots \zeta_k^{(r-1)} \zeta_1^{(r)} \cdots \zeta_k^{(r)}$$

labeled w^r such that $\zeta^{(i)} = \zeta_1^{(i)} \cdots \zeta_k^{(i)} \in \mathcal{L}^{-1}(w)$ for all $i = 1, \ldots, r$ and $\zeta_k^{(r)} \neq \eta_k$. There exist i, j with $1 \leq i < j \leq r$ such that $\zeta^{(i)} = \zeta^{(j)}$. If $\zeta^{(j)}$ starts at I, then, since W is a prefix, we get $\zeta_k^{(r)} = \eta_k$, which is a contradiction. So $\zeta^{(i)}$ does not start at I. Then $\zeta^{(i)} \cdots \zeta^{(j-1)}$ is a cycle labeled w^{j-i} that does not start at I. This is impossible due to Lemma 3.1. Thus $d_\varphi(w^r, |w^r|) = 1$.

(2) Let $u \in \mathcal{I}_s \cap \mathcal{S}(W^*)$ and $\mathsf{T}(u)$ denote the set of all vertices J in G such that there is a path τ labeled w terminating at J. If u is not a synchronizing word for W, then $\mathsf{T}(u) = \{J_1, \ldots, J_d\}$ for some $d \geq 2$, where $J_1 = I$. There is a path π in G terminating at J_2 such that $\mathcal{L}(\pi) = u$. Take the shortest path from I to the initial vertex of π, say α, and let $p = \mathcal{L}(\alpha) \in \mathcal{P}(W)$. Consider $pu = \mathcal{L}(\alpha\pi) \in \mathcal{P}(W^*)$. If $v, w \in W^*$, then $vpu, uw \in \mathcal{F}(W^*)$. Since $u \in \mathcal{I}_s$, we have $vpuw \in \mathcal{F}(W^*)$. So $pu \in W^e$; i.e., $pu \in W^*$. Since W is a prefix, we have $J_2 = I$, which is a contradiction. So u is a synchronizing word for W. Thus every word in $\mathcal{I}_s \cap \mathcal{S}(W^*)$ is a synchronizing word for W. □

Corollary 4.7. *Let $W \in \mathcal{W}_\mathcal{A}$ be a prefix and X_W be a shift of finite type. Put $\varphi = \varphi_W$. Then the following are equivalent.*

(i) *W is X_W-indecomposable.*
(ii) *There is $r \in \mathbb{N}$ such that $d_\varphi(w^r, |w^r|) = 1$ for all $w \in \mathcal{I}_s \cap W^+$.*
(iii) *Every word in $\mathcal{I}_s \cap W^+$ is a synchronizing word for W.*

Proof. Let $\alpha \in \mathcal{R} \setminus W$ so that $\alpha^k \in W$ for some $k \geq 2$. Since X_W is a shift of finite type, there is $N \geq 1$ such that $\alpha^l \in \mathcal{I}_s$ for all $l \geq N$. Then $\alpha^{kN} \in \mathcal{I}_s \cap W^+$ but α^{kN} is not a synchronizing word for W. Thus (iii) implies (i). Also, letting $w = \alpha^{kN}$, we have $d_\varphi(w^r, |w|^r) \geq 2$ for all $r \in \mathbb{N}$. Therefore (ii) implies (i). This completes the proof. □

Corollary 4.8. *Let $W \in \mathcal{W}_\mathcal{A}$ be a bifix. Then the following are equivalent.*

(i) *W is X_W-indecomposable.*
(ii) *There is a synchronizing word for W.*
(iii) *There is an r-synchronizing word for W.*
(iv) *The degree of φ_W is one.*

Proof. It follows from Proposition 4.6 that (i) implies (ii). Since W is a suffix, a symmetric argument can show that (i) implies (iii). Next, (ii) implies (iv), since W is a suffix. Similarly, (iii) implies (iv), since W is a prefix.

To see that (iv) implies (i), let $\mathcal{G} = (G, \mathcal{L})$ be the graph presentation for W with the central vertex I and $d_\varphi = 1$ where $\varphi = \mathcal{L}_\infty$. Then there exist $w = w_1 \cdots w_k \in \mathcal{F}(W^*)$ and $1 \leq i \leq k$ such that $d_\varphi(w, i) = 1$. That is, there is an edge e in G such that, whenever $\tau = \tau_1 \cdots \tau_k$ is a path in G labeled w, then $\tau_i = e$. If $v, u \in \mathcal{F}(W^*)$ and $vwu \in \mathcal{F}(W^*)$, then $d_\varphi(vwu, |v| + i) = 1$. So we may assume that $w \in W^*$. Let $\alpha \in W^e$. Since W is a suffix, we have $w\alpha \in M_s$. Hence there is a path $\pi = \pi_1 \cdots \pi_k\tau$ terminating at I with $\mathcal{L}(\pi) = w\alpha$ and $\mathcal{L}(\tau) = \alpha$. Then $\pi_i = e$, since $d_\varphi(w\alpha, i) = 1$. So there is a path from the initial vertex of e to I labeled $w_i \cdots w_k\alpha$; i.e., $w\alpha \in W^*$. Since W is a prefix, we have $\alpha \in W^*$. Thus W is X_W-indecomposable. □

Example 4.1. Let $W = \{0, 10, 1110, 1^40, 1^6\}$. Then W is a prefix and $\mathcal{P}_\circ = W \cup \{111\}$. So W is not X_W-indecomposable. Note that, whenever $m \in \mathbb{N}$ and $m \equiv 2 \pmod 3$, then $01^m0 \notin \mathcal{F}(W^*)$. Let $\bar{w} = 1^n$ for some $n \geq 1$. Then $n = 3p + q$, where $p \geq 0$ and $0 \leq q < 3$. Put $u = 01^{3-q}$ and $v = 110$. Then $u\bar{w}, \bar{w}v \in \mathcal{F}(W^*)$ but

$$u\bar{w}v = 01^{3-q}1^{3p+q}110 = 01^{3p+5}0 \notin \mathcal{F}(W^*);$$

hence $\bar{w} \notin \mathcal{I}_s$. So if $w \in \mathcal{I}_s \cap \mathcal{S}(W^*)$, then $0 \in \mathcal{F}(w)$. Since 0 is a synchronizing word for W, so is w. Thus every word in $\mathcal{I}_s \cap \mathcal{S}(W^*)$ is a synchronizing word for W. Also, if $w \in \mathcal{I}_s \cap W^+$, then $d_\varphi(w^2, |w^2|) = 1$. Thus the converses of Proposition 4.6 do not hold. Next, one can see that $\mathcal{R} = \{0, 10, 111\}$ (after simplification). Note that there is no r-synchronizing word for $\mathcal{R}$, while $\mathcal{R}$ is $\mathsf{X}_\mathcal{R}$-indecomposable. Thus the converse of Proposition 4.5 does not hold.

5. Cyclic codes and maximal monoids

Let $W \in \mathcal{W}_\mathcal{A}$ be cyclic and $\mathcal{G} = (G, \mathcal{L})$ the graph presentation for W. Put $\varphi = \mathcal{L}_\infty$. Since φ is a conjugacy, there exist $l, r \geq 0$ with the following property. For any $u \in \mathcal{A}^{l+}$ and $v \in \mathcal{A}^{(r+1)+}$ with $uv \in \mathcal{F}(W^*)$, there is an edge e in G such that, whenever $\pi = \tau\pi$ is a path in G labeled uv with $\mathcal{L}(\tau) = u$ and $\mathcal{L}(\pi) = v$, then τe is a path in G and e is the first edge of π. In other words, φ^{-1} has memory l and anticipation r. The number l is called a memory for W and r is called an anticipation for W. Let $u \in \mathcal{F}(W^*)$ and $|u| \geq l + r + 1$. If $u \notin \mathcal{I}_s$, then there exist $v, w \in \mathcal{A}^*$ such that $vu, uw \in \mathcal{F}(W^*)$ and $vuw \notin \mathcal{F}(W^*)$. Hence there exist two paths $\tau = \tau_1 \cdots \tau_{|u|}$ and $\pi = \pi_1 \cdots \pi_{|u|}$ in $\mathcal{L}^{-1}(u)$ such that the terminal vertex of τ_l does not coincide with that of π_l. This is impossible. Thus $u \in \mathcal{I}_s$. Therefore $\mathcal{F}(W^*) \cap \mathcal{A}^{(l+r+1)+} \subseteq \mathcal{I}_s$.

Lemma 5.1. *Let $W \in \mathcal{W}_\mathcal{A}$ be cyclic. Then $M_s \cap M_p = W^*$ and $\mathcal{S}_\circ \cap \mathcal{P}_\circ = W$. Also, $|\mathcal{P}_\circ| < \infty$ and $|\mathcal{S}_\circ| < \infty$. Hence M_p and M_s are finitely generated.*

Proof. Let $l \geq 0$ be a memory for W. If $\alpha \in M_s \cap M_p$, then $\alpha^{l+1} \in M_s \cap M_p$, so $\alpha^{l+1} \in W^*$. Hence $\alpha \in W^*$. Thus $M_s \cap M_p = W^*$. Let $w \in W$ and $w = up$ for some $u \in W^+$ and $p \in \mathcal{P}_\circ$. Since l is a memory for W, we have $pv \in W^*$ for any $v \in W^* \cap \mathcal{A}^{l+}$. Then $wv = u(pv) \in W^*$. Since W is a code, we have $p \in W^*$, which is a contradiction. Thus $w \in \mathcal{P}_\circ$. Similarly, $w \in \mathcal{S}_\circ$. Therefore $W \subseteq \mathcal{S}_\circ \cap \mathcal{P}_\circ$. Note that $\mathcal{S}_\circ \cap \mathcal{P}_\circ \subseteq M_s \cap M_p = W^*$. It follows that $\mathcal{S}_\circ \cap \mathcal{P}_\circ \subseteq W$. Thus $\mathcal{S}_\circ \cap \mathcal{P}_\circ = W$.

Next, let $\alpha \in M_p \setminus \{\varepsilon\}$. There is $p \in \mathcal{P}_\circ$ such that $\alpha \in W^* \cdot p$. Let $\alpha = up = vq$, where $u, v \in W^*$ and $p, q \in \mathcal{P}_\circ$ with $|p| < |q|$. Then $q = rp$ for some $r \in \mathcal{S}(u) \setminus \{\varepsilon\}$, so $u = vr$. Since l is a memory for W, we have $pw, qw \in W^*$ for any $w \in W^* \cap \mathcal{A}^{l+}$. It follows that $vr(pw) = u(pw) = v(qw)$. Since W is a code, we get $r \in W^*$, so $q \in W^+ \cdot p$. This is a contradiction. Thus there is a unique $p \in \mathcal{P}_\circ$ such that $\alpha \in W^* \cdot p$. Similarly, given $\beta \in M_s \setminus \{\varepsilon\}$, there is a unique $s \in \mathcal{S}_\circ$ such that $\beta \in s \cdot W^*$. Now, let $t = \max\{|w| | w \in W\}$ and $\alpha \in M_p \cap \mathcal{A}^{(r+t)+}$, where r is an anticipation for W. Then $\alpha = up$, where $u \in W^+$ and $p \in \mathcal{P}(W^*) \cap \mathcal{A}^{r+}$. Since r is an anticipation for W, it follows that $pw \in \mathcal{P}(W^*)$ for all $w \in W^*$. So $p \in M_p$. Hence $\alpha \notin \mathcal{P}_\circ$. Thus $|\mathcal{P}_\circ| < \infty$. Similarly, $|\mathcal{S}_\circ| < \infty$. Therefore M_p and M_s are finitely generated. □

Corollary 5.2. *Let $W \in \mathcal{W}_\mathcal{A}$ be cyclic. Then the following statements hold.*

(1) *$M_s = W^*$ if and only if $M_p = W^d$.*
(2) *$M_p = W^*$ if and only if $M_s = W^d$.*
(3) *$M_s = M_p$ if and only if W is X_W-indecomposable.*

Proof. (1) Let $l \geq 0$ be a memory for W. Let $M_P = W^d$. Since $M_s \subseteq W^d$, it follows from Lemma 5.1 that $M_s = M_s \cap M_p = W^*$. Conversely, let $M_s = W^*$ and $\alpha \in W^e$. Choose any $w \in W^* \cap \mathcal{A}^{l+}$. Then $\alpha w \in M_s$, so $\alpha w \in W^*$. Hence $\alpha \in \mathcal{P}(W^*)$. Thus $W^e \subseteq \mathcal{P}(W^*)$. By Lemma 2.2, we have $M_p = W^e$ or $M_p = W^d$. □

Let $W \in \mathcal{W}_\mathcal{A}$ be cyclic. If either $M_s = W^*$ or $M_p = W^*$, then $M_s \cdot M_p = W^e$. It turns out that $W^e \setminus (M_s \cdot M_p)$ is always finite. Denote by $\mathcal{D}(W)$ the collection of all finite X_W-indecomposable sets V with $V^* \supseteq W^*$.

Theorem 5.3. *Let $W \in \mathcal{W}_\mathcal{A}$ be cyclic. Let $r \geq 0$ be an anticipation for W and $l \geq 0$ a memory for W. Define $q = \min\{|\mathcal{P}_b|, |\mathcal{S}_b|\}$, where*

$$\mathcal{P}_b = \{\alpha \in M_p \setminus W^* | |\alpha| < r + t\},$$
$$\mathcal{S}_b = \{\beta \in M_s \setminus W^* | |\beta| < l + t\}$$

(after simplification) and $t = \max\{|w| | w \in W\}$. Then $|\mathcal{D}(W)| \leq 2^q - q$.

Proof. We first show that

$$W^e \cap \mathcal{A}^{(l+r+t)+} \subseteq M_s \cdot M_p. \tag{5.1}$$

To see this, let $\pi \in W^e$ and $|\pi| \geq l + r + t$. Choose any $w \in W^* \cap \mathcal{A}^{r+}$. Since r is an anticipation for W, we have $w\pi \in M_p$. It follows from the proof of Lemma 5.1 that $\mathcal{P}_\circ \subseteq \mathcal{F}(\mathcal{A}^{r+t-1})$. So there exist $u \in W^*$, $\alpha \in \mathcal{P}_\circ$, and $\beta \in \mathcal{S}(w)$ such that $w\pi = u\alpha$ and $\pi = \beta\alpha$. Note that $|\beta| > l$. Also, $\beta \in \mathcal{S}(W^*) \cap \mathcal{P}(W^e)$. This implies that $\beta \in M_s$. Thus $\pi = \beta\alpha \in M_s \cdot M_p$. This verifies (5.1). Consequently, $W^e \setminus (M_s \cdot M_p)$ is a finite set.

Next, if $\mathcal{P}_b = \emptyset$, then $\mathcal{P}_\circ = W$, so $W^e = M_s$. Thus $\mathcal{D}(W) = \{\mathcal{S}_\circ\}$. Similarly, if $\mathcal{S}_b = \emptyset$, then $\mathcal{D}(W) = \{\mathcal{P}_\circ\}$. So we may assume that $1 \leq q = |\mathcal{P}_b| \leq |\mathcal{S}_b|$. Now, let $V, U \in \mathcal{D}(W)$ and $V^* \cap \mathcal{P}_b = U^* \cap \mathcal{P}_b$, or equivalently,

$$(V^* \cap M_p) \setminus \mathcal{A}^{(r+t)+} = (U^* \cap M_p) \setminus \mathcal{A}^{(r+t)+}. \tag{5.2}$$

We claim that $V = U$, or equivalently, $V^* = U^*$. To see this, let $\pi \in V^* \cap \mathcal{A}^{(l+r+2t)+}$. By (5.1), there exist $\beta \in \mathcal{S}(W^*)$, $u \in W^* \cap \mathcal{A}^{(l+1)+}$, and $\alpha \in M_p \cap \mathcal{A}^{r+}$ such that $\pi = \beta u\alpha$. We claim that $\alpha \in V^*$. To see this, let $w \in V^*$. Since $u\alpha \in \mathcal{S}(V^*)$, we have $u\alpha w \in \mathcal{F}(W^*)$. Since r is an anticipation for W and l is a memory for W, we get $\alpha w \in \mathcal{P}(W^*)$. So, if $v \in V^*$, then $v\alpha w \in \mathcal{F}(W^*)$. Hence $\alpha \in V^e$, i.e., $\alpha \in V^*$, as claimed. We may assume that $|\alpha| < r + t$. From (5.2), we get $\alpha \in U^* \cap M_p$. Thus, if $w \in U^*$, then $\alpha w \in U^*$ and hence $\pi w \in \mathcal{S}(U^*)$. This implies that $V^* \cdot U^* \subseteq \mathcal{S}(U^*)$. A symmetric argument shows that $U^* \cdot V^* \subseteq \mathcal{S}(V^*)$. So

$$U^* \cdot V^* \cdot U^* \subseteq \mathcal{S}(U^*).$$

Thus $V^* \subseteq U^e$ or $V^* \subseteq U^*$. Therefore $V^* = U^*$; that is, $V = U$.

As a result, for $\Gamma \subseteq \mathcal{P}_b$, there is at most one set V in $\mathcal{D}(W)$ with $V^* \cap \mathcal{P}_b = \Gamma$. So $|\mathcal{D}(W)| \leq 2^q$. Finally, we will show that, if $V \in \mathcal{D}(W)$, then $|V^* \cap \mathcal{P}_b| \neq 1$. To see this, let $V \in \mathcal{D}(W)$ and $V^* \cap \mathcal{P}_b = \{\alpha\}$. Since $\alpha^2 \in (V^* \cap M_p) \setminus W^*$, we have $|\alpha^2| \geq r + t$. By the same argument as above, there exist $u \in W^+$ and $p \in M_p \cap V^*$ such that $\alpha^2 = up$ and $r \leq |p| < r + t$. So $p \in V^* \cap \mathcal{P}_b$. But $p \neq \alpha$. This is a contradiction. Thus there is no $V \in \mathcal{D}(W)$ such that $|V^* \cap \mathcal{P}_b| = 1$. Therefore $|\mathcal{D}(W)| \leq 2^q - q$. □

Corollary 5.4. *Let $W \in \mathcal{W}_\mathcal{A}$ be cyclic. Let $r \geq 0$ be an anticipation for W and $l \geq 0$ a memory for W. Define*

$$s = |\{\pi \in \mathcal{F}(W^*) | |\pi| < k + t\}|,$$

where $k = \max\{l, r\}$ and $t = \max\{|w| | w \in W\}$. Then $|\mathcal{D}(W)| \leq (\sqrt{2})^s$.

The next result describes (possibly, the same) X_W-indecomposable sets that form a basis of $\mathcal{D}(W)$ for a cyclic code W.

Proposition 5.5. *Let $W \in \mathcal{W}_{\mathcal{A}}$ be cyclic. Then the following statements hold.*

(1) *There is a unique X_W-indecomposable set V_s such that $(V_s)^* \supseteq M_s$.*
(2) *There is a unique X_W-indecomposable set V_p such that $(V_p)^* \supseteq M_p$.*
(3) *There is a unique element in $\mathcal{D}(W)$ if and only if $V_s = V_p$.*

Proof. (2) Let $r \geq 0$ be an anticipation for W. Let $V = \mathcal{P}_\circ$ and $\pi \in W^e$. Choose any $w \in W^* \cap \mathcal{A}^{r+}$. Then $w\pi \in V^*$, so $\pi \in \mathcal{S}(V^*)$. Hence $V^e \subseteq W^e \subseteq \mathcal{S}(V^*)$. Lemma 2.2 implies that V^e is a unique maximal monoid in $\mathcal{F}(W^*)$ containing V^*, i.e., M_p. Since X_W is a shift of finite type, there is $V_p \in \mathcal{W}_{\mathcal{A}}$ such that $(V_p)^* = V^e$.

(3) Let $V_s = V_p$. Then

$$M_p \cdot M_s \subseteq (V_p)^* \cdot (V_p)^* = (V_p)^*.$$

From the proof of (2), one can see that $W^e \subseteq \mathcal{S}(M_p) \cap \mathcal{P}(M_s)$. Hence

$$W^e \cdot W^e \subseteq \mathcal{S}(M_p) \cdot \mathcal{P}(M_s) \subseteq \mathcal{F}(M_p \cdot M_s) \subseteq \mathcal{F}(W^*).$$

So W^e is a monoid. Thus there is a unique element in $\mathcal{D}(W)$. □

The following is an example of a cyclic code W such that $W^* \subsetneq M_p, M_s \subsetneq W^d$.

Example 5.1. Let $W = W_1 \cup W_2 \cup W_3 \cup W_4$, where

$W_1 = \{0, 10, 011, 1011\}$,
$W_2 = \{3, 32, 223, 2232\}$,
$W_3 = \{13, 132, 1223, 12232\}$,
$W_4 = \{02, 102, 0112, 10112\}$.

One can check that W is cyclic. Let $\alpha = 01$, $\beta = 23$, $\gamma = 101$, and $\delta = 232$. Then

$$\mathcal{P}_\circ = W \cup \{\alpha, \gamma\} \quad \text{and} \quad \mathcal{S}_\circ = W \cup \{\beta, \delta\}.$$

Moreover, $\mathcal{D}(W) = \{\mathcal{P}_\circ, \mathcal{S}_\circ\}$. Also, $W^d = M_s \cup M_p$ and $W^e = M_s \cdot M_p$.

The following is an example of a cyclic code W such that $W \subsetneq V_s = V_p$.

Example 5.2. Let $W = W_1 \cup \cdots \cup W_4 \cup \{12\}$, where $W_1, \ldots, W_4$ are given as in Example 5.1. One can check that W is cyclic. Let $\alpha = 01$, $\beta = 23$, $\gamma = 101$, and $\delta = 232$. Then

$$\mathcal{P}_\circ = W \cup \{\alpha, \gamma\} \quad \text{and} \quad \mathcal{S}_\circ = W \cup \{\beta, \delta\}.$$

Moreover, $\mathcal{S}_\circ \cup \mathcal{P}_\circ$ is X_W-indecomposable. So

$$V_s = V_p = \mathcal{S}_\circ \cup \mathcal{P}_\circ.$$

Thus $\mathcal{D}(W) = \{\mathcal{S}_\circ \cup \mathcal{P}_\circ\}$. Also, $W^d = W^e = (\mathcal{S}_\circ \cup \mathcal{P}_\circ)^* = M_s \cdot M_p$.

There is also an example of a cyclic code $W \in \mathcal{W}_{\mathcal{A}}$ for which $|\mathcal{D}(W)| \geq 3$. We will omit the description.

References

[1] M.P. Beal, D. Perrin, Complete codes in a soficshift, Lect. Notes in Comput. Sci. 3884 (2006) 127–136.
[2] M.P. Beal, D. Perrin, Codes and soficconstraints, Theoret. Comput. Sci. 340 (2005) 381–393.
[3] M.P. Beal, D. Perrin, Symbolic dynamics and finite automata, in: Handbook of Formal Languages, Vol. 2, Springer, Berlin, 1997, pp. 463–505.
[4] J. Berstel, D. Perrin, Theory of Codes, Academic Press, New York, 1985.
[5] F. Blanchard, G. Hansel, Systèmes codés, Theoret. Comput. Sci. 44 (1986) 17–49.
[6] C. de Felice, A. Restivo, Some results on finite maximal codes, Theoret. Inform. and Appl. 19 (1985) 383–403.
[7] A. de Luca, A. Restivo, On some properties of very pure codes, Theoret. Comput. Sci. 10 (1980) 157–170.
[8] A. de Luca, A. Restivo, Synchronization and maximality for very pure subsemigroups of a free semigroup, Lect. Notes in Comput. Sci. 74 (1979) 363–371.
[9] S. Hong, S. Shin, Cyclic renewal systems, Theoret. Comput. Sci. 410 (2009) 2675–2784.
[10] D. Lind, B. Marcus, An Introduction to Symbolic Dynamics and Coding, Cambridge Univ. Press, 1995.
[11] J. Néraud, Completing circular codes in regular submonoids, Theoret. Comput. Sci. 391 (2008) 90–98.
[12] J. Néraud, Completing prefix codes in submonoids, Theoret. Comput. Sci. 356 (2006) 245–254.
[13] J. Néraud, Completing a code in a regular submonoidof the free monoid, Lect. Notes in Comput. Sci. 3354 (2005) 281–291.
[14] J. Néraud, C. Selmi, A characterization of complete finite prefix codes in arbitrary submonoids of A^*, J. Aut. Lang. Comb. 9 (2004) 103–110.
[15] J. Néraud, C. Selmi, Free monoidtheory: Maximality and completeness in arbitrary submonoids, Int. J. Alg. and Comp. 13 (2003) 507–516.
[16] A. Restivo, A note on renewal systems, Theoret. Comput. Sci. 94 (1992) 367–371.
[17] A. Restivo, Codes and local constraints, Theoret. Comput. Sci. 72 (1990) 55–64.
[18] A. Restivo, On codes having no finite completions, Discrete Math. 17 (1977) 309–316.